ESIP Winter Meeting - 7 January 2020

Session Title: Bringing Science Data Uncertainty Down to Earth - Sub-orbital, In Situ, and Beyond

Attendees (name / email / agency/company/institution / level of interest (observer/participant):

1. David Moroni / david.f.moroni@jpl.nasa.gov Jet Propulsion Lab / participant
2. Leila Belabbassi / leila-belabbassi@tamu.edu / Texas A&M University/ Gulf of Mexico Ocean Observing System/ GCOOS (observer)
3. Jon Hobbs / jonathan.m.hobbs@jpl.nasa.gov / Jet Propulsion Lab / participant
4. Mikael Kuusela / mkuusela@andrew.cmu.edu / Carnegie Mellon University / participant
5. Ge Peng/ gpeng@ncsu.edu /NCICS, NCSU/participant (online)
6. Chung-Lin Shie, UMBC/JCET, NASA/GSFC
7. Wm Sonntag wsonntag@mit.edu wasonntag@gmail.com
8. Yaxing Wei / weiy@ornl.gov / ORNL DAAC
9. Mike Little / ESTO-AIST NASA
10. Ed Armstrong / JPL
11. Crista Straub, online
12. Jason Tallant, online
13. Zachary Robbins, online

Agenda:

1. "IQC Uncertainty White Paper Status Summary and Next Steps" - Presented by: David Moroni (15 minutes)
2. "Uncertainty quantification for in situ ocean data: The S-MODE sub-orbital campaign" - Presented by: Fred Bingham (15 minutes)
3. "Uncertainty Quantification for Spatio-Temporal Mapping of Argo Float Data" - Presented by Mikael Kuusela (20 minutes)
4. Panel Discussion (35 minutes)
5. Closing Comments (5 minutes)

Notes:

1. David Moroni, IQC Uncertainty White Paper, Now on Figshare
    - Focus on discovery of breadth of perspectives on uncertainty for Earth science data, including mathematical, programmatic, observational, user
    - Mathematical: emphasis on probability, identified use cases particularly in remote sensing
    - Programmatic: considered agency approaches as well as US law and international agreements
    - Observational: addressed calibration/validation and product development

- User: explored ways uncertainty information can be effectively or ineffectively consumed, examples such as ISO-19157, UncertML
  - Shared interest in approaches. Utility depends on effective communication of information related to uncertainty
  - Sub-orbital and in-situ have their own priorities and concerns

2. Fred Bingham (UNC Wilmington), S-MODE (Submesoscale Ocean Dynamics Experiment) sub-orbital campaign - 5-yrs $30 M
   - Newest technologies now allow us to observe ocean features at the kilometer scale, which is unprecedented.
   - The S-MODE campaign will help with testing the hypothesis that kilometer scale ocean eddies make important contributions to vertical exchange of climate and biological variables in the upper ocean. Fred is the data manager..
   - Has 4 different types of aircraft measurements.
   - Standard ship-based measurements will be retrieved through a variety of ocean cruises on research vessels.
   - Autonomous and drifting instruments will also be used.
   - Due to the diverse breadth and quantify of data, UQ/UC will be quite challenging.
   - Measuring the same variable in different ways
   - Surface drifters can be used to measure the Lagrangian tracks.
   - ADCP (acoustic doppler current profilers) provide a quasi-Eulerian measurement that measures the instantaneous velocity of a water column.
   - There has been discussion or statement on how uncertainty should be disseminated or communicated; at the moment, it's currently left up to the PI of each campaign; currently there is at least 12 PIs involved in this.
   - A typical uncertainty in a well calibrated CTD is 0.002 degrees Celsius; the only type of error that isn't well addressed with CTD is representation error, which is the error associated with the representation of temperature in space and time, such as the difference of the temperature surrounding the single point of observation.
   - Many quantities are not measured directly, but rather are computed; errors are therefore propagated through complex algorithms, which makes uncertainties difficult to assess.
   - Difficult to estimate uncertainty when there is no existing ground-truth.
   - Top-down approach: obtain uncertainty by considering the noise characteristics of the instrument and propagating errors.
   - Bottom-up: obtain uncertainty by comparing measurement to ground truth.
   - UQ is not yet fully ingrained in the culture of observational oceanography.
   - Question: Have you talked with the Data Management Working Group at Marshall; Answer: I've been working closely with Deborath Smith, who's helped review the 90-page DMP.
   - Recommendation from Ed Armstrong: would be good to consider imposing metadata and data file encoding standards for S-MODE and all future similar types of NASA-funded sub-orbital campaigns.

3. Mikael Kuusela
    - Currently 3856 ARGO floats circulating in the ocean.
    - Provides measurements at 3 by 3 degree spatial resolution every 10 days.
    - ARGO rely on gridded data products where the irregular and sparse data are interpolated statistically onto a dense, regular grid.
    - Most existing ARGO data products either do not provide uncertainty information or use statistically ad hoc techniques for UQ.
    - Key challenge: uncertainties of the interpolated maps depend on the covariance length scales that vary from one location to another.
    - Further complications: large size and non-Gaussianity of the data.
    - We have studied mapping ARGO data locally, in a moving-window fashion, which alleviates both the computational and the modeling challenges; also enables more accurate UQ, among other benefits.
    - We perform a “local Gaussian process regression on Argo data, focusing a small spatio-temporal window neighborhood around a given point in space and time.
    - Assuming a stationary Gaussian process within the window, we first estimate the covariance parameters of the process and then use the fitted model to interpolate the temperature at the given point in space and time.
    - This procedure is repeated for overlapping, moving windows.
    - Uncertainties in the data depend strongly on length scales.
    - Correlation of uncertainty along zonal length scales are observed, and show substantial variance as a function of depth.
    - Confidence levels vary as a function of interval length.
    - Key point: obtaining realistic uncertainty requires proper handling of the length scales and the non-Gaussianity of the data.
    - UQ for ocean heat content: requires knowing the predictive covariance between 2 different locations of the observable quantity, which cannot be easily extracted from the previous methodology.
    - In many cases, it’s not enough to only communicate the pointwise uncertainties, because the covariances provide more meaningful information to understand uncertainty in space-time.

Panel Discussion:

- Estimation of uncertainty for derived quantities, such as ocean vorticity, may be dependent on the magnitude of the quantity; another instance of nonstationarity
- Is IQC uncertainty effort looking at land cover? Not specifically at this point but authors may be looking at other use cases in the near future
- What are some possible topics for the follow-on to the white paper?
    - Additional discussion of in-situ and sub-orbital data products, possible use cases arising from today’s presentations
    - Land cover observations similarities or differences in uncertainty quantification from ocean observations

    - Move in the direction of developing recommendations, working with NetCDF-UQ (Ken Kehoe) or UncertML
- Try to link the practices that are conducted in the relevant use cases to the recommendations in the white paper to establish the linkage and identify potential gaps in what the community is doing
- Looking for more invited speakers for monthly telecons